DOCUMENT RESUME

ED 458 656 CS 510 670

AUTHOR Bunz, Ulla K.
TITLE Usability and Gratifications--Towards a Website Analysis Model.
PUB DATE 2001-11-00
NOTE 27p.; Paper presented at the Annual Meeting of the National Communication Association (87th, Atlanta, GA, November 1-4, 2001).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150) -- Tests/Questionnaires (160)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Comparative Analysis; Higher Education; Pilot Projects; Protocol Analysis; User Needs (Information); *User Satisfaction (Information); *World Wide Web
IDENTIFIERS Uses and Gratifications Research; *Web Sites

ABSTRACT

This paper discusses Web site usability issues. Specifically, it assumes that the usability of a Web site depends more on the perception of the user than on the objectively assessable usability criteria of the Web site. Two pilot studies, based on theoretical notions of uses and gratifications theory and similar theories, are presented. In the first study, experts evaluated three Web sites on the national park, Mesa Verde, in a more formal approach based on criteria defined in the literature. In the second study, non-experts (eight undergraduate students in a Basic Communication course) evaluated the same three Web sites in a more informal and personal approach, using concurrent, or "thinking aloud," verbal protocol methods. Results show that overall assessment of the Web sites differs between experts and non-experts. Specifically, overall the Web site assessed as worst by the experts was liked most by the non-experts. Cognitive and emotional needs as defined by uses and gratifications seemed to make more of a difference with regard to Web site use, and less with regard to Web site evaluation. Results from these studies provide the basis for a user-centered Web site analysis model that may make use of but not depend on usability criteria defined by the literature. (Contains 82 references, 7 Web sites used/quoted in footnotes, 4 notes, and 2 tables of data. The survey instrument is attached.) (Author/RS)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 458 656

Usability and Gratifications – Towards a Website Analysis Model

Ulla K. Bunz
University of Kansas

Presented at the 87th annual conference of the National Communication Association,
Atlanta, GA, USA
November, 2001

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☐ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Contact:
Ulla Bunz
Communication Studies
University of Kansas
1440 Jayhawk Blvd.
Lawrence, KS 66045
Email: ulla@ku.edu

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
U. K. Bunz
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Abstract

This paper discusses website usability issues. Specifically, it assumes that the usability of a website depends more on the perception of the user than on the objectively assessable usability criteria of the website. Two pilot studies, based on theoretical notions of uses and gratifications theory and similar theories, are presented. In the first study, experts evaluated three websites on the national park Mesa Verde in a more formal approach based on criteria defined in the literature. In the second study, non-experts evaluated the same three websites in a more informal and personal approach, using concurrent, or "thinking aloud," verbal protocol methods. Results show that overall assessment of the websites differs between experts and non-experts. Specifically, overall the website assessed as worst by the experts was liked most by the non-experts. Cognitive and emotional needs as defined by uses and gratifications seemed to make more of a difference with regard to website use, and less with regard to website evaluation. Results from these studies provide the basis for a user-centered website analysis model that may make use of but not depend on usability criteria defined by the literature.

CS 510 670

BEST COPY AVAILABLE

Table of Contents

Rationale 3
Literature 3
Uses and Gratifications Theory 3
Cognitive Needs 4
Affective Needs 5
Personal Integrative Needs 5
Social Integrative Needs 5
Dependency Theory 6
Notion of Social Stability 6
Cognitive Effects 6
Affective Effects 6
Behavioral Effects 6
Theory of Pleasurability 6
Psycho-Pleasure 7
Physio-Pleasure 7
Ideo-Pleasure 7
Socio-Pleasure 7
Application of Theories to Usability Studies 7
Problem of Created Needs 7
Problem of Intra-personal Processes 8
Usability Literature Overview 8
Methodology 9
Choice of Websites 9
Expert Evaluation 10
"Thinking aloud" Usability Testing 10
Non-Expert Evaluation 11
Results 12
Expert Evaluation 13
Survey Questionnaire 13
Evaluation Comments 14
Non-Expert Evaluation 15
Visual Cues 15
Navigation Issues 16
Content Evaluation 17
Cognitive and Emotional Needs 17
Overall Discussion 17
References 18
Appendix 24
Expert Instructions 24
Table 1 26
Table 2 26

Rationale

Ideally, every website has a clearly defined purpose. Ideally, users recognize this purpose as the message to be communicated, and, ideally, if they feel addressed by it, they use the website successfully. In real online life, though, miscommunication between websites' messages and websites' users is common.

The purpose of this paper is to examine the uses and gratifications notion of active-choice directed satisfaction of needs in relation to website evaluation. The two discussed empirical studies combined the concept of web usability with the basic assumptions of uses and gratifications to develop a website analysis model that serves to explain individual website evaluation in context. Specifically, the studies placed an emphasis on the website user as an active receiver of the communicative message. It sought to examine the way a user perceives the website and subsequently assesses or judges it as to its usability. These studies also considered the actual usability of a website according to literature criteria and expert assessment.

The literature (i.e. Krug 2000) refers to the difficulty of assessing and evaluating a website, mostly because this process is subjective, and internal. This notion is supported by social construction of technology studies and theories (i.e., Markus 1987; Pinch & Bijker 1987; Latour 1991) that agree that technologies become socially accepted not when they are "perfect" but when a sufficiently large group of people agree on their usefullness. Thus, a user's perception of a website provides interesting information in addition to double checking a website's elements to expert criteria. The methodology used assessed both the individual's perception of the website (specifically: its usability) and also acknowledged expert evaluation.

Some people might not think of website use as a communicative process. However, a website contains graphical and textual elements, just as a book does, and thus can be considered a communication medium. Incidentally, much web research has focused on frequency of use, type of activity performed, or effect of web use. Some research, such as in interpersonal communication, may have taken the web as a context in which communication takes place. However, in other areas, such as human factors research, researchers are aware of the interaction between websites as entities and their users. This is not a one-way stimulus-response interaction. Users can communicate back to the website in various ways, including leaving the website, which would be like refusing talk in verbal communication.

As technology evolves, so must the theories that describe them. The uses and gratifications theory, which is one of the conceptual frameworks applied to this study, has long been criticized for not explaining "why" people use a certain medium, or "how" a certain gratification would be provided by using a medium. These studies investigate a very similar question, placing website usability in the communication context. The methodology used seeks to get at exactly the "why" and "how" of using a website. This research thus, investigates communication processes at a very fundamental level, from an as of yet little researched perspective. It also is based on the implied hypothesis that personal factors influence website evaluation more than do usability factors.

Literature

Uses and Gratifications Theory

Uses and gratifications research can be traced to the 1950s, but did not really become a popular theory until the 1970s, especially through the landmark book edited by

Blumler and Katz (1974). In 1996, Lin called uses and gratifications one of the most influential theories in the field of communication. While there is still certain empirical lack to distinguish between certain components such as needs and motivations, the theory has now been developed, applied, critiqued, and has managed to integrate and be compared to other theories, especially those dealing with new technologies.

One year before the landmark book, Katz, Gurevitch and Haas (1973) already laid out many of the important assumptions of this theory. The authors stipulate that people choose media actively rather than be overpowered by them. The use of media, which can take one of multiple roles such as entertainment, information, influence, etc. (based on Laswell, 1948), is influenced by social role and psychological predisposition of the individual, the actor. The authors state that media needs must be viewed in a social context, because some of the needs fulfilled by the media (such as entertainment, information) can also be fulfilled by other means, such as family, friends, sleep, drugs, etc., and sometimes even better.

In Katz, Blumler and Gurevitch (1974) the authors then clearly define the basic assumptions of the theory: it is concerned with the social and psychological origins of needs which generate expectations of the mass media and other sources which lead to differential patterns of media exposure, resulting in need gratifications and other consequences, mostly unintended ones. Underlying this are the assumptions that audiences are active and goal-directed, audiences are empowered to link need with gratification choice, and that the media compete with other sources of need satisfaction. They acknowledge that gratification can be obtained from one of three sources, media content, exposure to media per se, and/or social context.

Despite, or maybe because of criticism examined in a later section of this paper, the theory continued to grow, and recent theoretical pieces can be found in Galloway and Meek (1981), A. M. Rubin (Rubin 1994), and Ruggiero (2000). Both Palmgreen (1984) and Williams et al (1985) called for more uses and gratifications research with new media. A response to this call are, among others, Walker and Bellamy's (1991) work on remote control use, Dimmick et al.'s (2000) study on email versus phone use, and various studies on television (i.e., Bryant & Zillman 1984; Finn 1997; Massey 1985). A growing body of literature employs a uses and gratifications approach to study the world wide web (i.e., Cochenor 1996; Ebersole 2000; Eighmey & McCord 1998; Hunter 2001; Mings 1997; Nordstrom 2001; van Waes 2000; Wyman et al. 1997). Equally, a growing body of literature examines motivations and satisfaction of web use, which is related to uses and gratifications world wide web research (i.e., Atkinson & Kydd 1997; D'Ambra & Rice 2001; Eighmey 1997; Gelderman 1998; Goodhue, Klein & March 2000; Lindroos 1997; Mathiesona & Keil 1998).

The original uses and gratifications theory distinguishes between four categories of audience needs, first laid out be Katz, Gurevitch and Haas (1973) and discussed in most mass communication theory textbooks, such as Tan (1985). As some of these needs will apply to the study of web usability, they are briefly examined in the following paragraphs.

Cognitive Needs. Cognitive needs result from the desire for information in an increasingly information rich society. People seek to understand and know about their environment to make sense of contexts.

Much use of the world wide web results from a desire to gratify cognitive needs. A large study with thousands of participants in numerous countries conducted by Georgia Tech (1998) showed that the majority of users surf the web for educational purposes (85%), or for the purpose of information gathering (57%). A recent example can also be found in the web use directly following the terrorist attack on the United States on September 11, 2001 (Rainie 2001). As the Internet ached under the amount of people logged on and sites broke down, people had to turn to other media such as TV.

Affective Needs. Affective needs are connected to emotional experiences and many people's intrinsic desire for pleasure, entertainment, and aesthetics. Many uses and gratifications studies have examined this need, foremost television studies. In Katz, Gurevitch and Haas' (1973) original article, television viewing was not described as an escapist activity, which changed by 1985, when Williams et al. considered the TV a medium of escapism.

In addition to information, web users also seek to gratify their affective needs on the world wide web. According to the Georgia Tech (1998) study, 61% of web users use the web for entertainment, and 51% use it for wasting time. Much game playing, music downloading, and chat room conversation results from a desire to fill affective, emotional, or entertainment needs, though some of this interaction is also fueled by social needs.

Personal Integrative Needs. Personal integrative needs derive from people's desire to appear credible, be perceived as confident, and have high self-esteem. These needs are closely related to an individual's value system.

A connection between personal integrative needs and the world wide web may not be apparent at first. However, connections exist and are important. Taking a macro perspecitive, each corporate website strives to convey credibility. On a micro level, personal homepages often reflect very personal attitudes, including diary excerpts, personal art, writing samples, and open affiliation with social organizations. All these are examples of people's need to overtly characterize themselves, display their inner workings and beliefs, and potentially strengthen their self-esteem by publishing what they consider an accomplishment. The level of technological sophistication with which this display takes place is not necessarily related to the importance the creation of the website has for each individual.

Social Integrative Needs. Social integrative needs are affiliation needs. People want to be part of a group, and want to be recognized as part of this group. People also feel the need to "reach out" to family, friends, or others they perceive to be in need of help.

Internet technology such as the world wide web and email provide numerous venues of gratifying social integrative needs. Contact can be made easily through messages and chat rooms. Membership can be displayed through logos or passwords. In western civilizations in general and in the technology dominated United States in specific, "being online" has almost become a social requirement. None-compliance can lead to excommunication (Fortner 1995). Recent research has shown that people use the web most to send email (93%), but after that, people use the web more to gratify

cognitive and affective needs than social needs (http://www.pewinternet.org/reports/chart.asp?img=6_internet_activities.jpg).

Dependency Theory

Dependency theory developed in a response to uses and gratifications theory (Ball-Rokeach & DeFleur 1976; DeFleur & Ball-Rokeach 1982). It is based on the notion that "audiences depend on media information to meet needs and attain goals" (Littlejohn 1992, 368). Here, media, audiences and society affect each other reciprocally. However, the more audiences depend on media, the more influenced they are by them.

Notion of Social Stability. The degree of dependence on the media depends not just on the individual alone. DeFleur and Ball-Rokeach (1989) discuss the ripple effect that change can exert on social systems. Especially in times of high uncertainty, or times when established beliefs and values are questioned, people tend to become more dependent on the media, because the stability of their social system is questioned. This can be observed in US American's use of mass media such as television directly following the terrorist attacks on September 11, 2001 (Rainie & Kalsnes 2001).

Cognitive Effects. Ball-Rokeach and DeFleur outline five types of cognitive effects, ambiguity resolution, attitude formation, agenda setting, expansion of belief system, and value clarification. All five apply to the world wide web. It is used to find information and thus, resolve equivocality and help attidue formation. Agenda setting and gate keeping are still the case, as not all information makes it onto webpages. People learn from webpages, and thus change their beliefs and values.

Affective Effects. Mass media such as the Internet (Morris & Ogan 1996) also have affective effects, which concerns people's feelings and emotions. States such as fear and happiness are often created or triggered by movies. States such as frustration or interest can easily be triggered through web use.

Behavioral Effects. Behavioral effects really derive from cognitive and affective effects, as internal processes have externally observable consequences. As with the social and personal integrative needs discussed under uses and gratifications theory, behavioral effects of the world wide web can influence behavior both online and offline. Expert users may use and evaluate the web differently than non-experts, and they may also derive different information or satisfaction from it that will impact offline life.

Theory of Pleasurability

Pleasurability is an important concept in the human-computer interaction literature. Pleasurability is defined as "not simply a property of a product but [...] the interaction between a product and a person" (Jordan 2000, p. 12). This notion is based on work by C. S. Lewis (Lewis 1960). Lewis defines need pleasures that can be generated or fulfilled by moving people from one state into another. Providing water to a thirsty person fulfills a need and provides pleasure. Taking an item from a person can create a need and negative pleasure in the case of anger or frustration.

The notion of pleasureability has been expanded on by Tiger (1992), who devised the four pleasure framework that is reviewed here, and that is used in usability research now. After defining the four pleasures, Tiger created a fictive person (Janet Peters) and analyzed her life and product consumption based on the four pleasures. As Jordan points out, pleasurability can be applied to usability research. The four-pleasure framework does not explain why people experience pleasure, but simply what kind of pleasure they experience (Jordan 2000, 15). Website usability analysis can assess the "why."

Psycho-Pleasure. Similar to cognitive needs and effects, psycho pleasure is related to cognitive elements providing pleasure. Technologies that are easy to learn or use can provide psycho-pleasure. Products that are challenging without being frustrating, such as computer games, can provide psycho pleasure. On the web, this may relate to finding information one is seeking.

Physio-Pleasure. Physio pleasure is derived from stimulation and satisfaction of sensory organs. This may include touch, smell, or visual aesthetics. The smell of a new car, or the easy grip on a razor may provide physio pleasure. On the web, physio pleasure will be provided through the layout of a website, the color use, and the overall feeling of positively after leaving a website.

Ideo-Pleasure. Ideo pleasures are value pleasures. A person valuing the art will obtain art products. A person concerned with the environment will buy environmentally save products. On the web, ideo pleasure may work in reverse. Information that will violate people's values and deprive them of ideo pleasure will be avoided. On the other hand, people who are proud of their stock portfolio may set the website portal of their browser to a financial website, displaying their pride and interest in this topic.

Socio-Pleasure. Socio pleasure is very similar to the social integration needs discussed above. People derive pleasure from affiliating with others, connecting, sharing, and being associated with certain groups and their values. Many car commercials suggest if only one owned the car, one would be the focal point of an envious crowd. Online, "members only" areas can provide a feeling of affiliation. The following matrix visualizes the similarities of the notions discussed.

[Insert Table 1 about here.]

Application of Theories to Usability Studies

In using the theories above for the assessment of website usability, several concerns, considerations and limitations apply. The two most important ones are discussed here.

Problem of Created Needs. Rosengren (1974) admits that gratifications are difficult to measure, partially because of social norms of what certain media are good for. This could influence self-report data or even selection of media. Regarding media choice, Blumler (1979) clarifies that the idea of the active audience is not a yes/no distinction, but ranges along a continuum. In addition, some media invite active participation more

than others. People may use the same medium for different reasons. Blumler says that the seeking of gratifications should be seen as an environmental deprivation.

In the research studies described here, media choice was restricted. Subjects were not able to choose their own medium, but were required to use the world wide web. They were even provided with certain websites they had to access, though future research will allow for a choice of topic and thus, place more emphasize on active choice.

In experimental research, certain conditions always have to be created artificially. With this research, the goal was to allow subjects to develop their own motivations and needs as much as possible despite limited media choice. Within each websites, subjects were still able to choose actively where they wanted to click, and what information they wanted to seek, or pay attention to, as is explained in more detail in the "Methodology" section. Thus, subjects were still able to develop needs defined by uses and gratifications literature and related theories, though the urgency of these needs may or may not have been very intense.

Problem of Intra-personal Processes. Lometti et al. (1977) review uses and gratifications and voice critique. One of their points is as to how exactly people may know which needs will be filled by what gratifications, and thus, how can they choose actively these gratifications? Indeed, many of the needs this theory addresses itself to are underlying and even unconscious, which Elliot (1974) in the landmark book already presented, as well as some of the points Swanson (1977) later picks up. Elliot also believes the methods used in early uses and gratifications studies created a certain reality and thus, people gave the answers they were expected to give, or thought they were expected to give.

This problem still exists, even with the new medium of the world wide web. Fact is that people are not always aware of what they want, or why they like something. Sometimes they even believe they want one thing, but end up paying attention to another. In website use this frequently occurs when people seek specific information, but then see a link or image that seems interesting enough to them to follow its lead off topic, or even off site.

The "thinking aloud" methodology discussed later and applied in the non-expert study described later seeks to address and possibly remedy this problem. "Thinking aloud" methods tap into the unconscious evaluation process. The website analysis model this research program is working towards will relate to internal process, but to a certain degree still depend on verbalization of just these by subjects, hoping for a "telling more than we know" effect as discussed by Nisbett and Wilson (1977).

Usability Literature Overview

The term "usability" was well chosen, because the concept behind it is just that, the ability to use. More specifically in this context, usability describes the ease of use a website provides. Usability is not a web specific term, but has been applied to the web for several years now. Both content and design of a website influence its usability. As Critchfield (1998) points out, a well designed website appears more credible, regardless of true credibility of the information provided. Yet, according to the literature, a poorly designed website mostly will not retain visitors long enough to persuade through content. Regarding text, poor design usually simply means "too much," which is why text in

general should be kept short and concise online (Morkes & Nielsen 1997). Terminology used is of great importance (Kukulska-Hulme 1999). Regarding design, one of the main criteria is that design needs to be intuitive (Binstock 1999). There is a great and growing body of literature providing tips and guidelines for designing websites with usability (i.e., Adler 1992; Dillon 1994; Gould & Lewis 1985; Instone 1997; Krug 2000; Nielsen 2000; Radosevich 1997; Rubin, J. 1994; Spool 1997). Much of the usability literature is also influenced by human-factors, interface design and human-computer interaction literature (i.e., Coe 1996; Galitz 1997; Hix & Schulman 1991; Lansdale & Ormerod 1994; Lindgaard 1992; Mayhew 1999; Schriver 1997). General literature on usability and its importance also can be found easily (i.e., Dorazio & Stovall 1997; Henry 1998; James 2001; Jordan 1998; Nielsen 1992b; Nielsen 1993).

Finally, a new body of literature is developing in the area of testing usability (i.e., Anderson 1999; Andre, Williges & Hartson 1999; Bevan & Maxleod 1994; Dickstein & Mills 2000; Ehrlich, Butler & Pernice 1994; Gray & Salzman 1998; Tamler 1998; Tang & Solomon 1998; van der Meij 1997; Virzi 1992; Wichansky 2000). Usability testing through "thinking aloud" methods is reviewed in the next section.

Methodology

Choice of Websites

In both pilot studies, the expert and the non-expert evaluation studies, three websites on the national park Mesa Verde in Colorado were used. These existing websites were downloaded completely using free shareware (WebCopier)[1] and uploaded again on the university server to protect against change of the websites throughout the research process, as recommended and performed by van Waes (2000).

There were four main considerations that led to the choice of these particular websites. First, any site chosen had to be readily downloadable and comparatively small in size. WebCopier downloads the entire domain, rather than one page or sub-folder of a domain. Thus, suitable websites could not be part of a larger domain or site, such as was the case, for example, with the Mesa Verde page linking off the US National Parks website. Websites also had to be reasonable contained within themselves, not offering too many external links that would take subjects off the site and onto the world wide web. On the other hand, sites could not be too large in number of pages, as subjects had to be able to form an overall opinion on them within ten minutes.

A second issue to be taken into account was that the three websites all had to present the same kind of information with the same spin. For example, a tour operator's website trying to sell vacation packages to Mesa Verde, and a historical, archeological site on Mesa Verde would not have presented the same message or kind of information. The goal was to find websites that present the same content, but look different.

Third, closing in on the actual content and subject of the website, a topic had to be chosen that would not be familiar to users. As a matter of fact, had any of the subjects been familiar with any of the three chosen websites, they would have been excluded from the study.

Finally, the choice of topic itself was a difficult task. It had to relate to subjects at least to a certain degree, so that there was a chance of inspiring motivation and needs that would guide movement throughout the websites. As the two studies described here are

[1] Available at http://www.maximumsoft.com

really pilot studies for a larger dissertation project that will use students as subjects, the overall theme of "spring travel" was chosen. Eventually, subjects will be able to choose among one of three categories, "Festival" (Mardi Gras), "Beach" (Spring Break at Panama City Beach), or "Outdoors" (Mesa Verde). To simplify the pilot studies, only the Mesa Verde websites were elected. This affected the results of the studies, as will be described later.

In conclusion, three Mesa Verde websites were chosen[2], subsequently referred to as Site One (http://www.mesa.verde.national-park.com/), Site Two (http://www.visitmesaverde.com/index.htm), and Site Three (http://www.mesaverdetours.com/).

Expert Evaluation

The websites were evaluated and ranked according to their usability level by "experts." The "experts" were recruited from members of the Association of Internet Researchers[3] by volunteering and providing expertise credentials. A message was posted to the association's listserve[4], briefly describing the purpose of this study, and volunteers were encouraged to contact the main researcher at her personal email address. Ten people volunteered, each of them receiving instructions (see Appendix) and the addresses of the websites in rotating order. At the time of writing, seven out of the ten experts had responded, which provides sufficient information for analysis, based on Lewis (1994), and Rubin (Rubin, J. 1994, p. 93): "For the purpose of conducting a less formal usability test, recent research has shown that four to five participants will expose 80 percent of the usability deficiencies of a product, and that this 80 percent will represent most of the major problems."

Experts evaluated each website using a brief 10-question survey, and by providing comments. The survey was constructed based on established usability questionnaires (Brooke 1996; Ravden & Johnson 1989), though the ten questions present a greatly reduced number of evaluation items, and some of them were slightly reworded to change a computer system usability evaluation instrument into a website evaluation instrument. Mostly, the ten questions served as guidelines for the experts while navigating the sites, as they were not given any specific tasks for their evaluation.

"Thinking aloud" Usability Testing

"Thinking aloud" methods, also referred to as concurrent protocols, have been used as qualitative methodologies for a number of years (i.e. Jørgensen 1989; Lewis 1982; Nielsen 1993; Nielsen 1994), especially to improve writing skills (i.e., Ransdell 1985). The method has been described by various sources (i.e., Airken & Mardegan 2000; Boren & Ramey 2000; Ericsson & Simon 1993; Nisbett & Wilson 1977; Rowe 1985), many of which are based on Ericsson and Simon's 1980 account and "thinking aloud" model. As Boren and Ramey (2000) point out, not all the literature followed Ericsson and Simon's strict guidelines. Boren and Ramey suggest to add a speech

---

[2] Please note that these are the original URL's for the websites. They may have expired or changed since the sites were downloaded for the purpose of this research.

[3] Association website: http://aoir.org

[4] Message was posted on October 8, 2001. To see the listserve archive, go to http://www.aoir.org/mailman/listinfo/air-l

communication approach to existing usability "thinking aloud" testing methods, which would allow for more interaction between the researcher and the subject. One point of disagreement between theoretical literature and applied research is the amount of prompting on the researcher's side. While the researcher should not lead or interrupt the subject, some subjects also develop uncertainty if they are not prompted enough. Drummond and Hopper (1993) and Condon (1986) even examine the use and meaning of "yeah" and "ok" respectively in this context.

Overall, the "thinking aloud" methodology has been evaluated as to its reliability and validity (i.e., Fidler 1983; Russo, Johnson & Stephens 1989; Smagorinsky 1989). Nielsen (1994) even provides guidelines for estimating the number of subjects needed for a "thinking aloud" test, just as Virzi (1990; 1992) and Lewis (1994) provided information on sufficient sample size in usability testing.

The "thinking aloud" method has also already been applied to web studies, and specifically to web usability studies (i.e., Boren 1999; Buur & Bagger 1999; Sienot 1997; van Waes 2000; Virzi, Scorce & Herbert 1993) and has been deemed an effective way of assessing website usability.

Regarding number of subjects and methodology, the literature draws some clear conclusions. Virzi, Sorce and Herbert (1993) compared three usability evaluation methods, the heuristic method that asks experts to evaluate the usability of a product or website, the thinking aloud method that asks "naïve" [their wording] subjects to do the same, and the performance method, where subjects are asked to perform tasks as fast as possible without evaluation or thinking aloud. Their conclusions include that experts are more efficient at determining minor problems than naïve users; and that the most efficient way of determining usability of a website is to combine heuristic and thinking aloud methods. My study follows this advice.

The literature also points to the low number of subjects required for successful and efficient usability testing. Repeatedly, authors point out that only four or five subjects are needed to determine about 80% of all usability problems (i.e., Nielsen 1994; Rubin J. 1994; Virzi 1992). Nielsen (1994) also reports that observing 36 usability experts, the average number of subjects used by each expert was 8.8 people. Virzi (1992) also shows that nine subjects are needed to find 95% of usability problems on average. However, Nielsen also points out that the final number of subjects should be influenced by the experience the experimenter has with usability testing in general, and thinking aloud methods in specific (Nielsen 1994, p. 393). Though experimenters don't have to be skilled specialists or very experienced to perform "thinking aloud" usability testing (Nielsen 1992a), the literature also shows that it doesn't hinder the research to have more subjects. The down sides are more time, more work, possibly higher financial costs, and smaller increments of new information. For example, Nielsen (1994, p. 390) shows that in his study the increase due to last subject between the first and the second subject is 20%, while the increase between fifth and sixth subject is only 5%.

Non-Expert Evaluation

A total of 8 subjects was drawn from a Basic Communication Course at a large mid-western university. Each subject met with the researcher one-on-one in the department's research Lab. Subjects were notified that the main purpose of the research

was to "find out what's going on inside of you while you use a website." A brief sample "thinking aloud" exercise was conducted to familiarize subjects with the procedure.

Then, subjects were asked to access three different web sites, one after the other. All three websites were about Mesa Verde, a national park in Colorado. The order in which these sites were accessed was rotated. While navigating the website, movement throughout the site was video recorded. Subjects were asked to "think aloud," which was audio recorded. Though not specifically requested, the comments gave insight to subjects' thought processes, the links they clicked on, and any opinions or emotions they formed, or experienced (such as slight frustration, excitement, curiosity, etc.) while using the website. Subjects used the websites for at least five minutes. Subjects were asked to move to the next website after about 10 minutes, even if they desired to continue using the website.

After navigating all three websites, subject were asked several questions, such as to compare the three websites according to visual impression, ease of navigation, content, an overall impression. Subjects were asked to explain their answers with examples from the websites. This required recall, but very immediate recall. Subjects were also asked specific questions regarding comments they made during the "think aloud" procedure, such as, "You said, 'Oh, this [an animated graphic of a book] is kind of neat. I'm clicking on it to see where it goes. [pause] Hm. [clicking again as nothing happens] That's weird.' Can you talk about that situation for a moment?" All subjects were also asked whether they would return to any of the websites, and if so, which one and why.

Results

On purpose, no formal requirements were provided that would identify "experts." In general, the literature (i.e. Virzi, Sorce & Herbert 1993) distinguishes experts and non-experts simply by the fact that experts have studied the topic of usability, and non-experts have not. The literature also points out that a problem with usability testing in general, and heuristic and verbal protocol analysis methods in specific is that many procedural differences exist that make it difficult to compare results (i.e. Bailey, Allan & Raielle 1992; Boren & Ramey 2000; Bowers & Snyder 1990; Deffner 1990). However, as the literature also points out that despite difficulties, and despite using different methods, experts and non-experts tend to identify the same most important usability flaws in systems and websites (Bailey, Allan & Raielle 1992; Virzi, Sorce & Herbert 1993), this research used both methods.

It is important to note that the main purposes of these studies were not to pin-point usability flaws that could be corrected to improve the evaluated websites. Instead, the main purpose of the expert study was to provide an evaluation of websites based on expert, literature defined, and formal usability criteria; and possibly to rank order the three websites according to their usability levels. The main purpose of the non-expert study was to provide an evaluation of websites based on non-expert, personal, and informal usability criteria. Obviously, personal opinion played into the expert evaluations in the open ended questions. Nonetheless, it can be assumed that these opinions are informed by the formal study of usability issues these experts have engaged in.

Finally, it is also important to note that the results reported below are preliminary, as analysis of all data has not been completed. The results reported here can, however, be

taken as strong indicators. Final analysis and results will be made available upon completion of the dissertation this research is part of.

Expert Evaluation

The seven self-announced experts participating in the expert evaluation study came with a variety of skills and usability evaluation experiences. Four were male, three were female. Three were Ph.D. students, one an undergraduate, and three were faculty members. All reported experience with usability on a theoretical level, as they had studied this topic. Three reported having engaged in usability testing before, two of which tested usability of computer systems and software, and one usability of political websites. Two reported having taught courses on usability and usability testing. Interestingly, the overall results of the expert evaluation were not always clear, and at some points even contradictory. This supports this research's implied hypothesis that personal factors influence website evaluation more than usability factors.

Survey Questionnaire. As sample size was very small (seven participants), only basic frequencies were run on the answers of the survey questions. However, results show beautifully how incoherent and inconsistent expert evaluations were. This does not reflect on the experts or their skills with evaluating usability. Instead, this is largely due to the little formalized nature of usability evaluation methods. It does question thus, the validity and reliability of any expert evaluation, as a different expert may have come to a completely different solution.

[Insert Table 2 about here.]

As can be seen from the means and sometimes multiple modes, experts did not agree on specific questions. For example, for question 3 (The information on the screen is easy to see and read) on Site One, the reported mean was a neutral 3.00, with an equal amount of experts disagreeing and agreeing. However, tendencies can be interpreted from these results. For example, though the overall ratings of all three sites are not statistically significantly different, experts tended to view Site Two as having highest usability of all three, followed by Site One, and Site Three having lowest usability, as was also the researcher's personal opinion. Possibly, with a higher sample size, these tendencies could have been statistically significant.

Careful interpretation also allows to draw other conclusions, for example with regard to color use. The usability literature states that color can be overused easily in website design. Overall, experts agreed that the use of color helped to make displays clear (Question 2) on Site Two, where only brown and black were used (mean 3.57; mode 5); but did not help on Site Three, where an abundance of text and background colors were used (mean 2.00; mode 1). Thus, these results correspond with usability criteria in the literature.

A last question to be discussed here is question 5 (It is easy to find information on this website). This question related both to content and to navigation issues. Site One had all links on the bottom of each page. Site Two had a left hand menu bar. Site Three was a frames page with the links located in the left-hand frame. Overall, Site One provided the largest amount of information in long paragraphs and pages over several screens. As can be seen in Table 2, Site Two received the highest mean rating (3.71) for this question,

with modes along the positive evaluation options. Site Two actually is the closest to standard information providing websites that users may be used to. Site One, which may have been most convenient through the left-hand links that never moved, received the lowest ranking (mean 2.29; mode 1). Despite the length of the pages, Site One actually received a slightly positive rating (mean 3.14; mode 2). This seems to indicate that frames do not have a positive effect on navigation, as a left hand navigation bar fulfills an equal purpose. This also indicates that shorter pages with fewer screens are preferred over longer pages. Both of these indicators correlate with usability criteria defined in the literature (Nielsen 2000).

Thus, a preliminary conclusion from the survey questionnaire states that experts do not agree statistically significant on their evaluation of websites, but in depth analysis reveals that their overall tendencies seem to align themselves with the usability criteria defined in the literature, as was expected.

Evaluation Comments. Experts were also able to make open ended comments below each website evaluation survey. In addition, some volunteered comments in emails. For example, expert 6 supported this research's implied hypothesis without knowing about it by stating, "I find usability testing very emotional and subjective. When I look at a site quickly, often the colours, tone of language and images will set the main attitude towards the site. If I initially respond positively to it, usability issues such as navigation and consistency are not as important." Quite obviously, this expert was speaking as a human user, not as a trained expert when typing these comments.

Though Site Two was called "slightly amateurish" in terms of graphic design by expert 7, it was described as "more professional" by expert 6, who also identified that this site does not "follow suggested HCI standards, such as "get to top" navigation, aids, clicking on the Mesa Verde picture (logo) to take you to the home page, etc." This was also noted by expert 4, who was "surprised" by this. On the other hand, expert 3 thought that this site had "perfectly adequate usability," though he believed that was "undercut by the poor user experience," as he found it difficult to elicit the main purpose of the site.

Navigation comments tended to be intertwined with comments regarding the amount of information provided on a site. For example, regarding Site One, expert 4 states, "The long pages confused me with all the information on one page. It's hard to navigate through different pages. I must learn how to use it before I can use it." This statement, having to learn before successful use, gets to the root of usability literature that states that websites with usability need to be intuitive (Binstock 1999). Expert 1 noted regarding navigation on Site One that "Navigation at bottom of page difficult – have to scroll all the time." Equally, regarding navigation on Site Three, she mentioned that "frames are a pain since horizontal scrolling is required." And expert 3 states that "although one can find the desired information, it takes serious work," mostly due to poor navigation and information overload issues. He noted the same problem of redundant and unnecessary yet incomplete information on Site One, as did expert 4. Expert 7 even mentioned terms from the usability literature in his comments regarding Site One. Specifically, he suggested the designers "should (re)-acquaint themselves with the 'inverted pyramid' principle" which states that more specific information should be located deeper within the site, and asked, "Ever heard about the 'keepit above the fold'

principle?" as in a low resolution, only header and image were displayed, requiring scrolling before any information could be obtained at all.

Almost all experts commented on use of color and animation one way or other, at least one noting positively the color schemes on Site One (expert 6), which were mostly standard text and link colors on a white/gray background. Specifically, experts tended to attack Site Three as "very busy…background adds to confusion…red text is hard to read," (expert 4); "too much animation and too much misuse of color" (expert 1); "text hard to read against background colors" (expert 6); simply "terrible!" (expert 3); or "What a dog! Normally, I'd have left a site like this after the first 5-10 sec, never to return!" (expert 7). Expert 3 mentioned the "tacky fire logo on the bottom of the page" on Site One, something those non-expert who found it noted as "fun" and generally positive. The only exception here was expert 2, who thought this site "more technically sophisticated without making any use of flashy graphics." This statement surprised the researcher in particular, as the top frame on Site Three has a green background, and big yellow dollar sign animated (turning) graphics.

Overall then, experts were quite verbal and very open and direct in their comments regarding the three websites. Interestingly, they tended to note negative aspects of websites, and only rarely pointed to positive aspects. Krug (2000) mocks the typical usability evaluator, who finds one thing he or she doesn't like, and generalizes this to all users based on the logic of "I'm a web user, and I don't like it," so others must also not like it. The experts in this usability evaluation test did not fall prey to that mistake, as they repeatedly referred to established usability criteria to support their criticism, and tended to use the proverb "I," rather than generalizing. Though these experts did not agree on all points, overall they provided exactly what was intended, a more formal expert evaluation based mostly on literature defined criteria.

Non-Expert Evaluation

Eight non-experts participated in the "thinking aloud" non-expert usability evaluation of the three websites. Six of these non-experts were women, two were men. No demographics were assessed. Subjects were instructed on the purpose and procedure of the study. Then, each practiced the "thinking aloud" method by using the university's website. They then looked at the three websites in rotating order. It is important to note that the results reported below are only preliminary results, as full transcription and analysis of audio-recorded "thinking aloud" sessions is not completed at time of writing. Much of the information presented here actually comes from the debriefing interviews and notes the researcher took during the "thinking aloud" tests. As with the expert evaluation above, results should be seen as tendencies.

Visual Cues. As reported above, experts greatly criticized the color schemes on Site Three as being too extreme, colorful, busy, or animated. Interestingly, at least five of the eight non-experts greatly preferred the color scheme of Site Three. It was considered less "boring," "more professional," and "more fun." Two subjects remarked that they did not like the colors on Site Three. When asked to explain further they mentioned the green background of the top frame, stating that they personally disliked green. To them this dislike seemed like enough explanation for rejecting the entire color scheme on this site. One subject remarked favorable upon the brown color used on Site Two, saying, "since

this is about an outdoorsy kind of place, it fits." Subjects did not agree on all color questions though. For example, one subject stated he did not like the red text, while another stated specifically that the red text was "neat." Obviously, personal preference plays an important part here.

For the most part, subjects also liked animation. They tended to note moving advertisements immediately without finding disturbance by them. One subject noted that the turning dollar signs on Site Three distracted him, but three subjects particularly liked them. Animation could also be found within each site: a burning camp fire at the bottom of one page in Site One, animated cartoon characters at the bottom of one page in Site Two, and page-turning books on one page in Site Three. All subjects saw the page-turning books and immediately felt compelled to click on these icons because of their movement. They continued to like the icons after nothing happened, though some reported slight frustration or annoyance because the link did not seem to work. [Note: This link links the page to itself and thus, nothing changes on the screen.] Reactions to the other animated icons was similar. Subjects were not disturbed by them. Typical reactions were, "Oh, cool!" followed by an attempt to click on the icon. These reactions are distinctly different from expert reactions reported above.

The last visual cues to be discussed here concern link colors. Nielsen (2000) recommends to use standard link colors, such as blue for a regular link, and purple for a visited link. These colors were used on Site One. Links on Site Two were black and underlined on the main pages, and brown and not underlined on the menu bar. The only exceptions were links on the menu bar identified with yellow stars. On Site Three, menu links were aqua, while links in the main pages were also standard color. Three subjects remarked on their inability to tell which link they had already visited. An example from Site Two: "Did I already go here? I think so. No. Yeah I did. [clicks on it] I think so. [clicks on back]. I've seen this page before." [Note: Actually, the subject had not been on the page before.] Likewise, three subjects specifically mentioned that they liked being able to tell which link they'd already visited. The fact that links were not underlined in the menu bar on Site Two did not seem to bother any subject. The fact that menu bar links did not change color bothered subjects on Site Two, but not on Site Three. Overall, subjects liked Site Three best visually.

Navigation Issues. Subjects navigated through the sites fairly easily, using the side menu bar or bottom links. Only one subject noticed that on Site Two, the side menu links were provided at the bottom of each page in small font. This subject remarked on this additional option favorable, stating that "Oh, that's good." When asked to expand on that during the interview he stated that he found it bothersome having to scroll back up to the top in order to click on a link, and this was more convenient. Interestingly, none of the subjects mentioned missing a "top" link that would allow upwards motion through the page.

Several subjects made comments about not finding what they expected when clicking on certain links. This was particularly the case with the "breaking news" link on Site Two. "I guess this is/would be news to them," two subjects echoed each other unknowingly when finding honors information on the "breaking news" page.

When confronted with long pages of text, subjects tended to scroll down these pages with increasing speed. Yet, only two subjects ever stopped scrolling before

reaching the bottom of the page before clicking on "back." This might have been due to the laboratory situation. Overall, subjects liked Site One or Three best for navigation.

Content Evaluation. Evaluation of content was difficult to assess for subjects. Those who were interested in the topic spent significantly longer on each site than those who were not. Those who were interested also tended to read some of the text, while others looked mostly at pictures. This supports the notion that websites have to relate to users to keep their attention and supports the decision to use three categories of sites for the main research that will follow these pilot studies. Overall, subjects felt overwhelmed by the amount of information displayed on all three sites. Several pointed to tables or brief paragraphs and made comments such as, "Yeah see, that's what I like." In debriefing interviews, several metioned that they surely would have read more about these things if only they had been interested. However, even those subjects who were initially interested in the topic suffered from information overload while often being unable to find specific information they were looking for. Five subjects made statements that if they were to write a paper or report on this topic, Site One would probably be most helpful.

Cognitive and Emotional Needs. Cognitive and emotional needs and motivations played into website use and evaluation in several ways. One subject in particular kept interrupting her own "thinking aloud" session, turning to the researcher and saying things such as, "If the site doesn't catch my interest right away, I move on." And "If it doesn't relate to me I go right past it." This subject also said in her debriefing interview, "If these sites had been about shopping, at least they would have been about something important." Clearly, the subject had no motivation other than the research situation that kept her moving through the websites.

However, especially those subjects who reported an interest in outdoor activities and hiking seemed to develop certain information needs and begin searching for answers throughout the website. One subject in particular began clicking through Site Three rapidly, muttering comments such as, "Hm. Naw, that's not it. Hm. Okay, no, oh, that's interesting. [upon noticing a link to an audio file of Hilary Clinton speaking on Mesa Verde] Hm." Asked later about this episode the subject reported that she had wanted to find out information about a certain trail that she remembered reading about on an earlier site she evaluated, but she couldn't find it. Clearly, goal oriented navigation differs from general browsing, which is supported by the literature.

Overall Discussion

The purpose of this research was to connect usability research and theories such as uses and gratifications to find out how and why users evaluate and use websites as they do. The studies presented above have to be analyzed much more in detail before certain conclusions can be drawn. However, it can be observed that in many instances, expert and non-expert evaluations of websites differed. This is particularly so in the overall evaluation of Site Three, which experts liked least, and overall, non-experts liked most.

Relevance of the website topic seems an important predictor or indicator to their movement throughout the site. Without interest, needs seem to mostly center around looking at pretty pictures and moving off the site fast. With interest, needs can be created by the content of the website.

References

1. Airken, L. M., & Mardegan, K. J. (2000). "Thinking aloud": data collection in the natural setting. Western Journal of Nursing Research, 22 (7), 841-853.
2. Atkinson, M., Kydd, C. (1997). Individual characteristics associated with world wise web use: An empirical study of playfulness and motivation. The DATA BASE for Advances in Information Systems, 28 (2), 53-61.
3. Bailey, R. W., Allan, R. W., & Raiello, P. (1992). Usability testing vs. heuristic evaluation: A head-to-head comparison. Proceedings of the Human Factors Society 36yh Annual Meeting, Monterey, 409-413.
4. Ball-Rokeach, S. J., & DeFleur, M. L. (1976). A dependency model of mass-media effects. Communication Research, 3, 3-21.
5. Blumler, J. G. (1979). The role of theory in uses and gratifications studies. Communication Research, 6, 9-36.
6. Blumler, J. G., & Katz, E. (Eds.) (1974). The uses of mass communications. Current perspectives on gratifications research. Beverly Hills, CA: Sage Publications.
7. Boren, M. T. (1999). Conducting verbal protocols in usability testing: Theory and practice. Ph.D. dissertation, Univ. Washington, Seattle, WA.
8. Boren, M. T., & Ramey, J. (2000). Thinking aloud: Reconciling theory and practice. IEEE Transactions on Professional Communication, 43 (3), 261-278.
9. Bowers, V., & Snyder, H. (1990). Concurrent versus retrospective verbal protocols for comparing window usability. Proceedings of the Human Factors Society, 34th Annual Meeting, Orlando, 1270-1274.
10. Brooke, J. (1996). SUS: A 'Quick and Dirty' usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & I. L. McClelland (Eds.), Usability evaluation in industry (pp. 189-194). London: Taylor & Francis.
11. Bryant, J., & Zillman, D. (1984). Using television to alleviate boredom and stress. Journal of Broadcasting, 28, 1-20.
12. Buur, J., & Bagger, K. (1999). Replacing usability testing with user dialogue. Communications of the ACM, 42 (5), 63-??.
13. Cochenor, J. J., et al. (1996). Image maps on the world wide web: The uses and limitations. Proceedings of Selected Research and Development Presentations at the 1996 National Convention of the Association for Educational Communications and Technology, Indianapolis, IN.
14. Condon, S. (1986). The discourse functions of OK. Semiotica, 60 (1/2), 73-101.
15. D'Ambra, J., & Rice R. E. (2001). Emerging factors in user evaluation of the world wide web. Information and Management, 38, 373-384.
16. Deffner, G. (1990). Verbal protocols as a research tool in human factors [symposium abstract]. Proceedings of the Human Factors Society, 34th Annual Meeting, Orlando, 1263-1264.
17. DeFleur, M. L., & Ball Rokeach, S. J., (1982). Theories of mass communicaton. New York: Longman.
18. DeFleur, M. L., & Ball-Rokeach, S. J. (1989). Theories of mass communication. New York: Longman.
19. Dimmick, J., Kline, S., & Stafford, L. (2000). The gratification niches of personal email and the telephone. Communication Research, 2 (2), 227-248.

20. Drummond, K., & Hopper, R. (1993). Some uses of *Yeah*. Research on Language and Social Interaction, 26 (2), 203-212.
21. Ebersole, S. (2000). Uses and gratifications of the web among students. Journal of Computer-Mediated Communication, 6 (1). Available online at http://www.ascusc.org/jcmc/vol6/issue1/ebersole.html
22. Eighmey, J. (1997). Profiling user responses to commercial web sites. Journal of Advertising Research, 37 (3), 59-66.
23. Eighmey, J., & McCord, L. (1998). Adding value in the information age: uses and gratifications of sites on the world wide web. Journal of Business Research, 41 (3), 187-194.
24. Elliot, P. (1974). Uses and gratifictions research: A critique and a sociological alternative. In J. G. Blumler and E. Katz, (Eds.), The uses of mass communications. Current perspectives on gratifications research. Beverly Hills, CA: Sage Publications.
25. Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. Psychological Review, 87 (3), 215-251.
26. Ericsson, K. A., & Simon, H. A. (1993). Protocol analysis: Verbal reports as data. Cambridge, MA: MIT Press.
27. Fidler, E. J. (1983). The reliability and validity of concurrent, retrospective, and interpretive verbal reports: an experimental study. In P. Humphreys, O. Svenson, & A. Vari (Eds.), Analysing and aiding decision processes, (pp. 429-440). Amsterdam: North-Holland.
28. Fortner, R. S. (1995). Excommunication in the information society. Critical Studies in Mass Communication, 12, 133-154.
29. Galloway, J. J., & Meek, F. L. (1981). Audience uses and gratifications: An expectancy model. Communication Research, 8, 435-449.
30. Gelderman, M. (1998). The relation between user satisfaction, usage of information systems and performance. Information and Management, 34 (1), 11-18.
31. Georgia Tech (1998). WWW user survey. Available online at http://www.cc.gatech.edu/gvu/user_surveys/
32. Goodhue, D., Klein, B., & March, S. (2000). User evaluations of IS as surrogates for objective performance. Information and Management, 38 (2), 87-101.
33. Hunter, C. D. (2001). The uses and gratifications of the world wide web. Available online at http://www.asc.upenn.edu/usr/chunter/webuses.html
34. Jordan, P. W. (2000). Designing pleasurable products. An introduction to the new human factors. London: Taylor & Francis.
35. Jørgensen, A. H. (1989). Using the thinking-aloud method in system development. In G. Salvendy and M. J. Smith (Eds.), Designing and using human-computer interfaces and knowledge based systems, (pp. 743-750). Amsterdam: Elsevier.
36. Katz, E., Blumler, J. G., & Gurevitch, M. (1974). Utilization of mass communication by the individual. In J. G. Blumler and E. Katz, (Eds.), The uses of mass communications. Current perspectives on gratifications research. Beverly Hills, CA: Sage Publications.
37. Katz, E., Gurevitch, M., & Haas, H. (1973). On the use of the mass media for important things. American Sociological Review, 38, 164-181.

38. Krug, S. (2000). Don't make me think!: a common sense approach to web usability. Indianapolis, IN: Que.
39. Laswell, H. D. (1948). The structure and function of communications in society. In L. Bryson (Ed.), *The communication of ideas*. New York: Harper.
40. Latour, B. (1991). Technology is society made durable. In J. Law (Ed.), A sociology of monsters: Essays on power, technology & domination, 103-131. London: Ruledge.
41. Lewis, C. (1982). Using the 'thinking-aloud' method in cognitive interface design. Research Report RC-9265, IBM T. J. Watson Research Center, Yorktown Heights, NY, USA.
42. Lewis, C. S. (1960). The four loves. London: Fount.
43. Lewis, J. R. (1994). Sample sizes for usability studies – Additional considerations. Human Factors 36 (2), 368-378.
44. Lin, C. A. (1996). Looking back: the contribution of Blumler and Katz's 'Uses of mass communication' to communication research. Journal of Broadcasting & Electronic Media, 40 (4), 574-581.
45. Lindroos, K. (1997). Use quality and the world wide web. Information and Software Technology, 39, 827-836.
46. Littlejohn, S. W. (1992). Theories of human communication (4th Edition). Belmont, CA: Wadsworth.
47. Lometti, G. E., Reeves, B., & Bybee, C. R. (1977). Investigating the assumptions of uses and gratifications research. Communication Research, 4, 321-338.
48. Markus, M. K. (1987). Toward a 'critical mass' theory of interactive media: Universal access, interdependence and diffusion. Communication Research, 14, 491-511.
49. Massey, K. B. (1995). Analyzing the uses and gratifications concept of audience activity with a qualitative approach: media encounters during the 1989 Loma Prieta earthquake disaster. Journal of Broadcasting & Electronic Media, 39, 328-342.
50. Mathiesona, K., & Keil, M. (1998). Beyond the interface: ease of use and task/technology fit. Information and management 34 (4), 221-230.
51. Mings, S. (1997). Uses and gratifications of online newspapers: A preliminary study. Electronic Journal of Communication, 7 (3).
52. Morris, M., & Ogan, C. (1996). The Internet as a mass medium. Journal of Communication, 46 (1)Journal of Communication, 46 (1), 39-50.
53. Nielsen, J. (1992a). Evaluating the thinking aloud technique for use by computer scientists. In H. R. Hartson and D. Hix (Eds.), Advances in human-computer interaction, Vol. 3, (pp. 69-82). Norwood, NJ: Ablex.
54. Nielsen, J. (1993). Usability Engineering. Boston, MA: AP Professional.
55. Nielsen, J. (1994). Estimating the number of subjects needed for a thinking aloud test. International Journal of Human-Computer Studies, 41, 385-397.
56. Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. Psychological Review, 84, 231-259.
57. Nordstrom, M. A. (2001). Assessment of web site effectiveness. Available online at http://personal.bgsu.edu/~nordstr/webassess.htm

58. Palmgreen, P. (1984). Uses and gratifications: A theoretical perspective. In R. N. Bostrom & B. H. Westley (Eds.). Communication Yearbook 8, (pp. 20-55). Newbury Park, CA: Sage.
59. Pinch, T. F. & Bijker, W. E. (1987). The social construction of facts and artifacts: Or how the sociology of science and the sociology of technology might benefit each other. In W. E. Bijker, T. P. Hughes & T. F. Pinch (Eds.) The social construction of technological systems: New directions in the sociology & history of technology (17-50). Cambridge, Massachusetts: MIT Press.
60. Rainie, L. (2001, September 15). How Americans used the Internet after the terrorist attack. Pew Internet and American life project, available online at http://www.pewinternet.org/reports/toc.asp?Report=45.
61. Rainie, L., & Kalsnes, B. (2001). The commons of the tragedy. Available online at http://www.pewinternet.org
62. Ransdell, S. (1995). Generating thinking-aloud protocols: impact on the narrative writing of college students. American Journal of Psychology, 108 (1), 89-98.
63. Ravden, S. J., & Johnson, G. I. (1989). Evaluating usability of human-computer interfaces: A practical method. Chichester: Ellis Horwood.
64. Rosengren, K. E. (1974). Uses and gratifications: A paradigm outlined. In J. G. Blumler and E. Katz, (Eds.), The uses of mass communications. Current perspectives on gratifications research. Beverly Hills, CA: Sage Publications.
65. Rowe, H. (1985). The "thinking-aloud" methodology. In R. J. Sternberg (Ed.), Problem solving and intelligence, (pp. 99-117). Hillsdale, NJ: Erlbaum.
66. Rubin, A. M. (1994). Media uses and effects: A uses-and-gratifications perspective. In J. Bryant and D. Zillmann (Eds.), Media effects: Advances in theory and research, (pp. 417-436). Hillsdale, NJ: Erlbaum.
67. Rubin, J. (1994). Handbook of usability testing: how to plan, design and conduct effective tests. New York: John Wiley & Sons.
68. Ruggiero, T. E. (2000). Uses and gratifications theory in the 21st century. Mass Communication and Society, 3 (1), 3-36.
69. Russo, J. E., Johnson, E. J., & Stephens, D. L. (1989). The validity of verbal protocols. Memory & Cognition, 17, 759-769.
70. Sienot, M. (1997). Pretesting web sites: a comparison between the think-aloud and the plus-minus method. Journal of Business and Technical Communication, 11, 469-483.
71. Smagorinsky, P. (1989). The reliability and validity of protocol analyses. Written Communication, 6, 463-479.
72. Swanson, D. L. (1977). The uses and misuses of uses and gratifications. Human Communication Research, 3, 214-221.
73. Tan, A. (1985). Mass communication theoris and research. New York: John Wiley & Sons.
74. Tiger, L. (1992). The pursuit of pleasure. Boston: Little, Brown & Company.
75. Van Waes, L. (2000). Thinking aloud as a method for testing the usability of websites: the influence of task variation on the evaluation of hypertext. IEEE Transactions on Professional Communication, 43 (3), 279-291.
76. Virzi, R. A. (1990). Streamlining the design process: Running fewer subjects. Proceedings of the Human Factors society 34th Annual Meeting, (pp. 291-294). Orlando, FL, 8-12 October.

77. Virzi, R. A. (1992). Refining the test phase of usability evaluation: how many subjects is enough? Human Factors, 34 (4), 457-468.
78. Virzi, R. A., Sorce, J. F., & Herbert, L. B. (1993). A comparison of three usability evaluation methods: heuristic, think-aloud, and performance testing. Proceedings of the Human Factors and Ergonomics Society, 37rd Annual Meeting (Santa Monica: HFES), 300-303.
79. Walker, J. R., & Bellamy, Jr., R. V. (1991). Gratifications of grazing: an exploratory study of remote control use. Journalism and Mass Communication Quarterly, 68 (3), 422-431.
80. Wichansky, A. M. (2000). Usability testing in 2000 and beyond. Ergonomics, 43 (7), 998-1006.
81. Williams, F., Phillips, A. F., & Lum, P. (1985). Gratifications associated with new communication technologies. In Rosengren, K. E., Wenner, L., & Palmgreen, P., Media gratifications research, pp. 241-254. Beverly Hills, CA: Sage.
82. Wyman, S. K., McClure, C. R., Beachboard, J. B., & Eschenfelder, K. R. (1997). User and system-based quality criteria for evaluating information resources and services available from federal websites: Final Report. Research report available from sponsoring agency, OCLC Online Comupter Library Center, Inc., Dublin, OH [BBB03546].

Websites used/quoted in footnotes

1. http://www.maximumsoft.com/ (To download the WebCopier)
2. http://www.pewinternet.org/reports/chart.asp?img=6_internet_activities.jpg (For an overview table of Amercian's Internet activities)
3. http://www.mesa.verde.national-park.com/ (Mesa Verde Site One used)
4. http://www.visitmesaverde.com/index.htm (Mesa Verde Site Two used)
5. http://www.mesaverdetours.com/ (Mesa Verde Site Three used)
6. http://aoir.org (Association of Internet Researchers' homepage)
7. http://www.aoir.org/mailman/listinfo/air-l (Association of Internet Researchers' listserve archive)

Appendix

Researcher address & information

Date

Dear participant,

Thank you for providing your expertise in evaluating the usability of websites.

Attached you will find three identical sheets that should be used for evaluating the websites. The only difference on the sheets are the URLs for three different websites, all providing information on Mesa Verde in Colorado.

You do not have to evaluate all three sites, but it would help me most if you did.

Recommended steps:

1. At the bottom of this sheet, please provide a brief explanation why you have expertise to evaluate usability of websites.
2. Turn to the first evaluation sheet and type in the first URL into your web browser.
3. Take no more than 10 minutes maximum to click through the website.
4. Apply your own personal evaluation criteria, and please pay attention to navigation, content and consistency issues.
5. Check off the 10 questions on the sheet according to the scale provided.
6. Feel free to add personal comments in the space provided at the bottom of each sheet.
7. On the scale provided (7-point Likert scale), please rank your overall evaluation of this website.

8. If you wish to help more and invest more time, please repeat this process (Step 2-7) for the second and third sheet/URL provided.

Thank you so much!
Please address all questions to **Researcher name and email**.

Name (optional):_____

Your experience/expertise with website evaluation and usability issues:

URL
Website address after uploading of saved websites

1 – strongly disagree 2 – disagree 3 – neutral 4 – agree 5 – strongly agree

	Please circle the appropriate number.					
1.	Information appears to be organized logically on the screen	1	2	3	4	5
2.	The use of color helps to make the displays clear	1	2	3	4	5
3.	The information on the screen is easy to see and read	1	2	3	4	5
4.	Overall, screens appear uncluttered	1	2	3	4	5
5.	It is easy to find desired information in this website	1	2	3	4	5
6.	Overall, the website is easy to use	1	2	3	4	5
7.	The method of selecting options (e.g. from a menu) is consistent throughout the website	1	2	3	4	5
8.	From the user's point of view, the content of the website is complete	1	2	3	4	5
9.	From the user's point of view, the website is well and clearly organized	1	2	3	4	5
10.	Overall, the website is inconsistent	1	2	3	4	5

Evaluation comments:

Overall, I would rate the usability of this website (please circle):

No usability			Neutral			Highest Usability
1	2	3	4	5	6	7

Note: The same sheet was provided two more times with the different website URLs. The order of the websites was rotated.

BEST COPY AVAILABLE

Table 1: Similarities between theory components

Uses and Gratifications Theory	Dependency Theory	Theory of Pleasurability
Cognitive needs	Cognitive effects	Psycho-pleasure
Affective needs	Affective effects	Physio-pleasure
Personal integrative needs	Behavioral effects	Ideo-pleasure
Social integrative needs		Socio-pleasure

Note: This matrix does not presume to equate the needs, effects and pleasures aligned horizontally. Instead, they are aligned to visualize similarities between the different approaches.

Table 2
Expert Evaluation of Websites – Basic Frequencies

	Site One		Site Two		Site Three	
	Mean	**Mode(s)**	**Mean**	**Mode**	**Mean**	**Mode**
Quest. 1	3.29	2, 3, 4	4.14	4	3.00	2, 3, 5
Quest. 2	2.71	2, 3, 4	3.57	5	2.00	1
Quest. 3	3.00	2, 4	4.00	4	2.14	1
Quest. 4	3.14	4	3.71	4	2.29	1
Quest. 5	3.14	2	3.71	3, 4, 5	2.29	1
Quest. 6	2.86	4	3.86	4	2.71	4
Quest. 7	3.29	2, 4	3.29	4	3.71	4
Quest. 8	3.29	3	3.43	4	2.29	2
Quest. 9	3.14	4	3.29	4	2.43	2
Quest. 10	2.14	2	2.14	2	3.29	3, 5
Overall	4.17	4, 6	4.83	4	2.8	2
	Note: Questions 1-10 were on a 5-point scale with 1-strongly disagree and 5-strongly agree; the "overall" question was on a 7-point scale with 1-no usability and 7-highest usability					
Legend: Quest. 1	Information appears to be organized logically on the screen					
Quest. 2	The use of color helps to make the displays clear					
Quest. 3	The information on the screen is easy to see and read					
Quest. 4	Overall, screens appear uncluttered					
Quest. 5	It is easy to find desired information in this website					
Quest. 6	Overall, the website is easy to use					
Quest. 7	The method of selecting options (e.g. from a menu) is consistent throughout the website					
Quest. 8	From the user's point of view, the content of the website is complete					
Quest. 9	**From the user's point of view, the website is well and clearly organized**					
Quest. 10	Overall, the website is inconsistent					
Overall	Overall, I would rate this site					